

Appl. No.: 09/943,579
Amdt. dated April 1, 2004
Reply to Office Action of January 16, 2004

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently amended) A computer~~Computer~~ apparatus for gene prediction comprising:

a processor that is adapted to execute instructions that cause the processor to

create a plurality of units for that predict predicting-gene locations in a subject genomic sequence, wherein each unit of said units is capable of providing respective intermediate indications of gene locations output by the unit; and

create a combiner coupled to each unit to receive the that receives said respective intermediate output indications of predicted gene locations, the combiner being formed of a comprising a Bayesian network which combines the predicted that combines said intermediate indications of gene locations using probabilities of gene locations of the subject genomic sequence[.]] and forming to form a final combined output for indicating predicted genes-gene locations in of the subject genomic sequence.

2. (Currently amended) The Computer-computer apparatus as claimed in claim 1 wherein the plurality of units is a plurality of expert systems.

3. (Currently amended) The Computer-computer apparatus as claimed in claim 1 wherein the Bayesian network accounts for includes probabilistic dependencies between individual units and dependencies between adjacent parts of the subject genomic sequence.

Appl. No.: 09/943,579
Amdt. dated April 1, 2004
Reply to Office Action of January 16, 2004

4. (Currently amended) ~~The Computer~~ computer apparatus as claimed in claim 3 wherein the Bayesian network combines the predicted gene locations according to

$$Y^* = \max_{Y_t} P(Y_t | E_1, \dots, E_n, Y_{t-1}^*) \\ E_i \in \{E, I\}$$

where t is location in the subject genomic sequence and E_1, \dots, E_n are the respective predictions (E for exon or I for intron ~~or intergenic region~~) made by individual units 1 through n, n being the number of units in the plurality.

5. (Currently amended) ~~The Computer~~ computer apparatus as claimed in claim 1 wherein the subject genomic sequence is a DNA or RNA sequence.

6. (Currently amended) ~~The Computer~~ computer apparatus as claimed in claim 1 wherein gene locations include exon predictions.

7. (Withdrawn) Computer apparatus as claimed in claim 6 wherein gene locations further include exon and intron predictions; and the final combined output indicates exons and introns of the predicted genes of the subject genomic sequence.

8. (Currently amended) ~~The Computer~~ computer apparatus as claimed in claim 1 wherein the Bayesian network comprises a table or set of probabilities of a given sub-sequence being a protein encoding exon prepared by applying training data to the computer apparatus, wherein said training data comprises character strings representing ~~is trained on~~ known genes of a known genome sequence.

9. (Currently amended) A method for automated gene prediction, comprising the steps of:

obtaining from a plurality of expert systems a plurality of respective preliminary gene location predictions for a subject gene in a subject genomic

Appl. No.: 09/943,579
 Amdt. dated April 1, 2004
 Reply to Office Action of January 16, 2004

~~sequence using a plurality of expert systems, preliminarily predicting gene locations in a subject genomic sequence;~~

~~inputting into a digital processor programmed to contain a Bayesian network a plurality of respective datasets representing said gene location predictions;~~

~~combining said respective datasets in said Bayesian network using a Bayesian network, combining the preliminarily predicted gene locations; and~~

~~to form forming from said Bayesian network a final combined data output indicating an improved predicted genes location for said subject gene in of the subject genomic sequence, wherein the Bayesian network includes probabilistic dependencies between individual expert systems and dependencies between adjacent parts of the subject genomic sequence accounting for dependencies between individual expert systems and dependencies between adjacent parts of the subject genomic sequence.~~

10. (Currently amended) ~~[[A]]The~~ method as claimed in claim 9 wherein ~~the step of combining using a~~ said Bayesian network combines ~~said plurality of datasets according to the probability equation~~

$$Y^* = \max_{Y_t} P(Y_t | E_1, \dots, E_n, Y_{t-1}^*)$$

$$E_i \in \{E, I\}$$

~~where wherein~~ ~~t is location of a subject gene~~ in the subject genomic sequence and E_1, \dots, E_n are the respective predictions (E for exon or I for intron or intergenic region) made by individual expert systems, n being the number of expert systems in the plurality.

11. (Currently amended) ~~[[A]]The~~ method as claimed in claim 9 wherein the subject genomic sequence is a DNA or RNA sequence.

12. (Currently amended) ~~[[A]]The~~ method as claimed in claim 9 wherein ~~said obtaining and said inputting each comprise predicting an exon location gene locations include exon predictions.~~

Appl. No.: 09/943,579
Amdt. dated April 1, 2004
Reply to Office Action of January 16, 2004

13. (Withdrawn) A method as claimed in claim 12 wherein gene locations further include exon and intron predictions; and the final combined output indicates exons and introns of the predicted genes of the subject genomic sequence.

14. (New) The method of claim 9 wherein said forming comprises creating from said Bayesian network a data output indicating an optimal combined predicted location for each said subject gene in the subject genomic sequence.

15. (New) The method of claim 9 comprising, prior to said combining, providing a training dataset to said processor representing a plurality of known gene locations on a known genomic sequence.

16. (New) The method of claim 15 comprising, constructing a conditional probability function or table from said training dataset.

17. (New) The method of claim 9 comprising selecting a subject gene that is unknown to said expert systems.

18. (New) A computer apparatus comprising means for performing the method of claim 9.